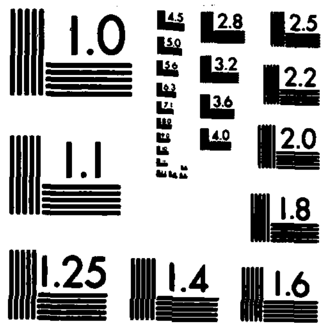
AD-A145 609 ON THE RATE OF MEAN CONVERGENCE OF FINITE LINEAR PREDICTORS OF MULTIVARIA..(U) NORTH CAROLINA UNIV AT CHAPEL HILL CENTER FOR STOCHASTIC PREC.. M POURAHMADI 1/1

UNCLASSIFIED JUL 84 TR-69 AFOSR-TR-84-0825 F/G 12/1 NL

END


1.0
1.1
1.25
1.4
1.6
1.8
2.0
2.2
2.5
2.8
3.2
3.6
4.0
4.5
5.0

CENTER FOR STOCHASTIC PROCESSES

Department of Statistics
University of North Carolina
Chapel Hill, North Carolina

ON THE RATE OF MEAN CONVERGENCE OF FINITE LINEAR PREDICTORS OF MULTIVARIATE STATIONARY STOCHASTIC PROCESSES

by

Mohsen Pourahmadi

Technical Report #69

July 1984

REPORT DOCUMENTATION PAGE

REPORT SECURITY CLASSIFICATION: UNCLASSIFIED	1b. RESTRICTIVE MARKINGS
SECURITY CLASSIFICATION AUTHORITY	3. DISTRIBUTION AVAILABILITY OF REPORT: Approved for public release; distribution unlimited.
DECLASSIFICATION DOWNGRADING SCHEDULE	
PERFORMING ORGANIZATION REPORT NUMBER(S): TR #69	5. MONITORING ORGANIZATION REPORT NUMBER(S): AFOSR-TR- 8 5
NAME OF PERFORMING ORGANIZATION: University of North Carolina; 6b. OFFICE SYMBOL (If applicable)	7a. NAME OF MONITORING ORGANIZATION: Air Force Office of Scientific Research
ADDRESS (City, State, and ZIP Code): Department of Statistics, Chapel Hill NC 27514	7b. ADDRESS (City, State, and ZIP Code): Directorate of Mathematical & Information Sciences, AFOSR, Bolling AFB DC 20332
NAME OF FUNDING/SPONSORING ORGANIZATION: AFOSR; 8b. OFFICE SYMBOL (If applicable): NM	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER: F49620-82-C-0009
ADDRESS (City, State, and ZIP Code): Bolling AFB DC 20332	10. SOURCE OF FUNDING NUMBERS

PROGRAM ELEMENT NO.	PROJECT NO.	TASK NO.	WORK UNIT ACCESSION NO.
61102F	2304	A5	

TITLE (Include Security Classification): ON THE RATE OF MEAN CONVERGENCE OF FINITE LINEAR PREDICTORS OF MULTIVARIATE STATIONARY STOCHASTIC PROCESSES

PERSONAL AUTHOR(S): Mohsen Pourahmadi

TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (Year, Month, Day)	15. PAGE COUNT
Technical	FROM ____ TO ____	July 1984	19

SUPPLEMENTARY NOTATION

COSATI CODES		
FIELD	GROUP	SUB-GROUP

18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number): q-variate stationary processes; spectral density matrix; linear least squares predictor; prediction error; exponential rate of convergence.

ABSTRACT (Continue on reverse if necessary and identify by block number)

Consider a multivariate weakly stationary stochastic process $\{X_n\}$ with the spectral density matrix f satisfying the boundedness condition. It is shown that if the entries of f are analytic functions of θ on $[-\pi,\pi]$, then the rate of convergence of the one-step ahead linear least squares predictor of $\{X_n\}$ based on a finite segment of the past, and the partial sum of the infinite linear least squares predictor of the process to the Kolmogorov-Wiener predictor is at least exponential.

20. DISTRIBUTION AVAILABILITY OF ABSTRACT: ☒ UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT ☐ DTIC USERS	21. ABSTRACT SECURITY CLASSIFICATION: UNCLASSIFIED
22a. NAME OF RESPONSIBLE INDIVIDUAL: CPT Brian W. Woodruff	22b. TELEPHONE (Include Area Code): (202) 767- 5027; 22c. OFFICE SYMBOL: NM

DD FORM 1473, 84 MAR — 83 APR edition may be used until exhausted. All other editions are obsolete.

SECURITY CLASSIFICATION OF THIS PAGE: UNCLASSIFIED

-A-

ON THE RATE OF MEAN CONVERGENCE OF FINITE LINEAR PREDICTORS OF MULTIVARIATE STATIONARY STOCHASTIC PROCESSES[1]

by

Mohsen Pourahmadi[2]
Center for Stochastic Processes
Department of Statistics
University of North Carolina
Chapel Hill, North Carolina 27514

This documents

theta [-pi,pi], sub n

Abstract

Considers a multivariate weakly stationary stochastic process $\{X_n\}$ with the spectral density matrix f satisfying the boundedness condition. It is shown that if the entries of f are analytic functions of θ on $[-\pi,\pi]$, then the rate of convergence of the one-step ahead linear least squares predictor of $\{X_n\}$ based on a finite segment of the past, and the partial sum of the infinite linear least squares predictor of the process to the Kolmogorov-Wiener predictor is at least exponential.

AMS 1980 Subject Classification : Primary 62M10; Secondary 60G12.

Keywords and Phrases: q-variate stationary processes, spectral density matrix, linear least squares predictor, prediction error, exponential rate of convergence.

[1]Research supported by the NSF Grant MCS-83012 40 and by the U.S. Air Force Office of Scientific Research Grant No. F 49620 82 C 0009.

[2]On leave from Department of Mathematical Sciences, Northern Illinois University.

1. Introduction.

Let $\{X_n\}_{n=-\infty}^{+\infty}$ be a q-variate ($1 \leq q < \infty$) weakly stationary stochastic process with the spectral density matrix f. In the Kolmogorov-Wiener theory of prediction, one assumes that the spectral density f and the infinite past $P = \{X_m;\ m \leq 0\}$ are known and the problem is to express the linear least squares predictor of X_1, denoted by $\hat{X}_1$, in terms of the components of f and P. It can be shown that $\hat{X}_1$ can formally be written as

$$\hat{X}_1 \sim \sum_{m=0}^{\infty} A_m X_{-m},$$

where $A_m = -C_0 D_{m+1}$, for the notation see Section 2.

In practice, however, one seldom knows f and P, and therefore it is necessary to approximate $\hat{X}_1$ by some properly chosen approximants based upon the available information. Very often, this available information is a finite segment of the past, i.e. $P_n = \{X_{-m};\ 0 \leq m \leq n-1\}$. Here, we consider the as yet impractical case of knowing f and P_n, and this will provide some useful results for the case when only P_n is known and f has to be estimated.

Now, knowing f and P_n there are two natural approximants for $\hat{X}_1$. The first one being the partial sum of the series representation of $\hat{X}_1 : \hat{X}_{1,n} = \sum_{m=0}^{n-1} A_m X_{-m}$. The second approximant is the linear least squares predictor of X_1 based on P_n, denoting this by $X^*_{1,n}$, we have $X^*_{1,n} = \sum_{m=0}^{n-1} A_{nm} X_{-m}$, where the coefficients A_{nm}, $0 \leq m \leq n-1$, can be obtained as the result of solving a minimization problem. Having identified these two approximants, it is natural to ask as how good they are. Namely it is of interest to know how fast $\|\hat{X}_1 - \hat{X}_{1,n}\|$ and $\|\hat{X}_1 - X^*_{1,n}\|$ approach zero as $n \to \infty$. The main purpose of this paper is to find rates of convergence to zero for these quantities, or the errors of approximating the infinite predictor $\hat{X}_1$ by finite predictors $\hat{X}_{1,n}$ and $X^*_{1,n}$.

It is obvious that if the series representation of $\hat{X}_1$ does not converge in the mean, then $||\hat{X}_1-\hat{X}_{1,n}|| = ||\sum_{m=n+1}^{\infty} A_m X_{-m}||$ does not go to zero at all as $n\to\infty$. This may suggest the type of (smoothness) conditions which one may need to impose on $\{X_n\}$ or f in order to obtain a fast rate of convergence.

The rate of convergence of $||\hat{X}_1-X^*_{1,n}||$ is dominated by that of $||\sum_{m=n+1}^{\infty} A_m X_{-m}||$ and $||\hat{X}_{1,n}-X^*_{1,n}||$, cf. (2.9). The rate of convergence of $||\hat{X}_{1,n}-X^*_{1,n}||$ is expected to play fundamental roles in some statistical problems related to the estimation of f, when only P_n is available. For the univariate processes this is demonstrated in [2,6], see also Section 4.

Let G be the qxq matrix of one-step ahead prediction error. Define $\sigma=\mathrm{tr}G$, $\sigma_0=(\mathrm{tr}G^{\frac{1}{2}})^2$ and $\sigma_n=||X_1-X^*_{1,n}||^2$. Note that for q=1, $\sigma=\sigma_0$. With this notation the outline of the paper is as follows. Section 2 is devoted to introducing some preliminary results establishing some useful relationships among σ_n,σ and σ_0.

For a univariate stationary process an important result of Grenander and Rosenblatt [4, Thm.4], [5, p. 188] gives necessary and sufficient conditions on f for the rate of convergence of $\sigma_n-\sigma_0(=\sigma_n-\sigma)$ to decrease at least exponentially to zero. The rate of convergence of $\sigma_n-\sigma_0$ plays an important role in finding the rate of convergence for finite linear predictors. The entire Section 3 is devoted to finding an appropriate multivariate extension of this result of Grenander and Rosenblatt, i.e. finding necessary and sufficient conditions on the matrix f such that $\delta_n=q\ \sigma_n-\sigma_0$ decreases to zero at least exponentially, cf. Theorems 3.2 and 3.7. The proof presented in [4,5] depends heavily on the analytical properties of orthogonal polynomials with respect to f. Since matricial orthogonal polynomials may not enjoy all of these properties or they are hard to prove, special attempt is made to prove the results of Section 3 without adherence to the orthogonal polynomials. This is done by adapting a multivariate version of a method used in Devinatz [3], cf. Theorem 3.1. Devinatz's proof of the necessity part of the Result of Grenander and Rosenblatt is still based on properties of orthogonal polynomials. We get around this by using a lemma due to Miamee [9], cf. Theorems 3.6 and 3.7.

In Section 4 under the boundedness condition on the spectral density, cf. (4.1), we show that the approximation errors; $\|\hat{X}_1-\hat{X}_{1,n}\|$, $\|\hat{X}_1-X^*_{1,n}\|$ and $\|\hat{X}_{1,n}-X^*_{1,n}\|$ have the same order as $\delta_n = q\ \sigma_n - \sigma_0$. Therefore, combining this with the multivariate extension of the result of Grenander and Rosenblatt we get exponential rate of convergence to zero for these errors provided that the entries of f are analytic functions on $[-\pi,\pi]$, cf. Theorems 4.1 and 4.2. Theorem 4.3 provides a rate of convergence for the finite linear interpolator of $\{X_n\}$ when only one value, say, X_0 is unknown.

2. Preliminaries

Let $(\Omega,\mathcal{F},P)$ be a probability space and $\mathcal{M}=L_0^2(\Omega,\mathcal{F},P)$ the Hilbert space of all complex-valued random variables on Ω with zero expectation and finite variance. The inner product in $\mathcal{M}$ is given by

$$(x,y)=Ex\bar{y}\ ,\ x,y\in M.$$

Following [8] for an integer $q\geq 1$, $\mathcal{M}^q$ denotes the cartesian product of M with itself q times, i.e. the set of all column vectors $X=(x_1,x_2,\ldots,x_q)$ with $x_i \in M$, $i=1,2,\ldots,q$. M^q is endowed with a Gramian structure. For X and Y in M^q their Gramian is defined to be the qxq matrix $(X,Y) = [(x_i,y_j)]_{i,j=1}^q$. It is well-known that M^q is a Hilbert space under the inner product $((X,Y)) = \text{trace}\ (X,Y) = \sum_{j=1}^q (x_j,y_j)$ and norm $\|X\| = \sqrt{((X,X))}$, provided the linear combinations are formed with constant matrices as their coefficients.

Throughout this paper, for a qxq matrix $A=(a_{ij})$, $trA=\sum_{i=1}^q a_{ii}$, $A^*=(\bar{a}_{ji})$, det A stands for the determinant of A, A^{-1} for the inverse of A when it exists and $\|A\|_E$ for the Eucleadian norm of A i.e. $\|A\|_E=(tr\ AA^*)^{1/2}$. For two qxq matrices A and B,

$A \geq B$ means that $A-B \geq 0$, i.e. $A-B$ is a non-negative definite matrix. The $q \times q$ identity matrix is denoted by I. Functions will be defined on $[-\pi,\pi]$ and we identify this interval with the unit circle in the complex plane in the natural way. Values of a function f defined on the unit circle will be denoted by $f(\theta)$ instead of $f(e^{i\theta})$. dm stands for the normalized Lebesgue measure on $[-\pi,\pi]$, i.e. $dm(\theta)=d\theta/2\pi$. For $1 \leq p \leq \infty$, $L^p(H^p)$ denotes the usual Lebesgue (Hardy) space of functions on the unit circle. $L^p_{q \times q}$ $(H^p_{q \times q})$ denotes the space of all $q \times q$ matrix-valued functions whose entries are in $L^p(H^p)$.

In the following we introduce a few concepts which are needed in this study. A more complete discussion with proofs can be found in [8,15].

Let $\{X_n;\ n \in Z\} \subset M^q$ be a q-variate weakly stationary stochastic process (WSSP) with the spectral density matrix f. The time domain of $\{X_n\}$ is defined by $\overline{sp}\{X_n: n \in Z\} = M(X)$, where $\overline{sp}\{\quad\}$ stands for the closed linear span of elements of $\{\quad\}$ in the metric of M^q. The spectral domain corresponding to the spectral density matrix f is denoted by $L^2(f)$ and is defined by $L^2(f) = \{\Phi;\Phi$ is a $q \times q$ matrix valued function with $\|\Phi\|^2_f = \int tr\, \Phi(\theta) f(\theta) \Phi^*(\theta) dm(\theta) < \infty\}$. It is known that $L^2(f)$ with inner product given by $((\Phi,\Psi))_f = \int tr\, \Phi f \Psi^* dm(\theta)$ is a Hilbert space and it is isomorphic to $M(X)$.

The problem of linear least squares prediction of a q-variate WSSP $\{X_n\}$ can be stated as the problem of finding the $q \times q$ matrices A_m such that the linear least squares predictor of X_1 based on X_m, $m \leq 0$, denoted by $\hat{X}_1$, can be written as $\hat{X}_1 = \underset{n\to\infty}{\text{l.i.m.}} \sum_{m=0}^{n} A_m X_{-m}$, where l.i.m. stands for the convergence in the norm of M^q (or in the mean). Let $G=(X_1-\hat{X}_1, X_1-\hat{X}_1)$ be the one-step ahead prediction error matrix. The process $\{X_n\}$ is said to be of <u>full-rank</u> if the matrix G is nonsingular.

From now on, we assume that the process $\{X_n\}$ is purely non-deterministic and full-rank. This is equivalent to assuming that $\{X_n\}$ has a spectral density f with $\log \det f \in L^1$ and $f=\Phi\Phi^*$, where $\Phi \in H^2_{q \times q}$ is an outer function with constant term $\Phi_0 = G^{1/2}$. We refer to Φ as the <u>generating function</u> of the process $\{X_n\}$. Since $\{X_n\}$ is full-rank, we can define a process $\{Y_n\}$ called the <u>normalized</u>

innovation process of $\{X_n\}$ such that $(Y_m, Y_n) = \delta_{m,n} I$. By using the Wold's decomposition [8] we have

$$X_n = \sum_{m=0}^{\infty} C_m Y_{n-m} \quad , \quad (C_0 = \Phi_0 = G^{\frac{1}{2}})$$

(2.1)

$$\Phi(\theta) = \sum_{m=0}^{\infty} C_m e^{im\theta}$$

Since Φ is analytic with no zeros in the open unit disc, hence Φ^{-1} is analytic and denoting its Fourier (Taylor) coefficients by D_k, $k=0,1,2,\ldots$, we have

$$\Phi^{-1}(\theta) = \sum_{m=0}^{\infty} D_m e^{im\theta}. \tag{2.2}$$

with this notation the linear least squares predictor of X_1 can formally be written as

$$\hat{X}_1 \sim -C_0 \sum_{m=0}^{\infty} D_{m+1} X_{-m} \quad , \tag{2.3}$$

where the infinite series on the right hand side of (2.3) may not converge in the norm of M^q.

We say that the linear least squares predictor of X_1 has a mean convergent autoregressive representation if the series in (2.3) converges in the norm of M^q.

In practice, however, we have only a finite number of observations $X_0, X_{-1}, \ldots, X_{-(n-1)}$ from which one has to construct a linear predictor for X_1. For this one has two options. The first one being the partial sum of the series in (2.3). Thus, with $A_m = -C_0 D_{m+1}$ and denoting this linear predictor by $\hat{X}_{1,n}$ we have

$$\hat{X}_{1,n} = \sum_{m=0}^{n-1} A_m X_{-m} \tag{2.4}$$

If $\hat{X}_1$ has a mean convergent autoregressive representation, then the norm of the error of approximating $\hat{X}_1$ by $\hat{X}_{1,n}$ is given by

$$\|\hat{X}_1 - \hat{X}_{1,n}\| = \|\sum_{m=n+1}^{\infty} A_m X_{-m}\| \to 0 \text{ as } n \to \infty. \tag{2.5}$$

A different linear predictor for X_1, denoted by $X^*_{1,n}$, can be obtained as the linear least squares predictor of X_1 based on n observations X_{-m}, $0 \le m \le n-1$, from the past. Thus, we have

(2.6) $$X^*_{1,n}=\sum_{m=0}^{n-1} A_{nm}X_{-m} ,$$

where A_{nm}, $0\leq m\leq n-1$, can be found by minimizing the norm $\|X_1-\sum_{m=0}^{n-1} B_m X_{-m}\|^2 =$ $\int tr(I-\sum_{m=1}^{n} B_m e^{im\theta})\, f(I-\sum_{m=1}^{n} B_m e^{im\theta})^*\, dm(\theta)$ over all choices of the constant matrices B_m.

Let us denote the minimum value of this norm by σ_n. Then, it can be shown that this minimum is attained for the polynomial $P_n(\theta)=\sigma_n \mathcal{V}_n(\theta)$, where $\mathcal{V}_n(\theta)=\sum_{m=0}^{n} \mathcal{V}_{nm}e^{im\theta}$ is the unique polynomial of degree n satisfying

(2.7) $$\int tr\, \mathcal{V}_n(\theta)f(\theta)\, e^{-ik\theta}dm(\theta)=\delta_{k,0} , \quad 0\leq k\leq n.$$

This unique minimizing polynomial $\mathcal{V}_n$ plays an important role in this paper. In the following we give a different characterization of the matrix-valued polynomial $\mathcal{V}_n$.

<u>Lemma 2.1</u>. Let $f=\Phi\Phi^*$ and $\sigma_0=(tr\ \Phi_0)^2$. Then,

(a) $\min \int tr(I-P\Phi)(I-P\Phi)^*dm= \int \|I-\sqrt{\sigma_0}\ \mathcal{V}_n\Phi\|^2_E\ dm,$

where the min is taken over all nth degree polynomials

$$P(\theta)=\sum_{m=0}^{n} P_{nm}e^{im\theta}.$$

(b) $\int \|I-\sqrt{\sigma_0}\ \mathcal{V}_n\ \Phi\|^2_E\ dm=q- \sigma_0/\sigma_n.$

<u>Proof</u>. (a) Let $M_n=\overline{sp}\{e^{ik\theta}\ \Phi;\ 0\leq k\leq n\}$ in $L^2(f)$. Then the minimum in (a) is attained for some H in M_n, if and only if $I-H\Phi$ is orthogonal to M_n or equivalently

$$\int tr(I-H\Phi)\Phi^*e^{-ik\theta}dm=0, \quad 0\leq k\leq n.$$

It is easy to show that $H=\sqrt{\sigma_0}\ \mathcal{V}_n$ satisfies the above relation. (b) follows from direct computation and the fact that $\int tr\ \mathcal{V}_n\Phi\ dm=\sqrt{\sigma_0}/\sigma_n$ and (2.7) with k=0.

Q.E.D.

Since Φ^{-1} is the isomorph of Y_0 (the normalized innovation) the following corollary, which follows immediately from Lemma 2.1, provides a different characterization of $\mathcal{U}_n$ in that $\sqrt{\sigma_0}\,\mathcal{U}_n$ is the isomorph of the linear least squares estimate of Y_0 based on the finite past. Note that with this interpretation, $q-\sigma_0/\sigma_n$ is the norm of the error of this estimator.

Corollary 2.2. Under the notation of Lemma 2.1, we have

$$\min \int tr(\Phi^{-1}-\sqrt{\sigma_0}P)f(\Phi^{-1}-\sqrt{\sigma_0}P)dm=\int tr(\Phi^{-1}-\sqrt{\sigma_0}\,\mathcal{U}_n)f(\Phi^{-1}-\sqrt{\sigma_0}\,\mathcal{U}_n)^*dm$$
$$=q-\sigma_0/\sigma_n.$$

It turns out that the rate of convergence of $q-\sigma_0/\sigma_n$ as $n\to\infty$ plays a crucial role in studying the rate of convergence of finite linear predictors and interpolator of a q-variate WSSP $\{X_n\}$. This is shown in the next two sections. In the following we record several inequalities which are useful in connecting rates of convergence of different quantities related to the problem studied here.

The following chain of inequalities follows from the definition of $\hat{X}_1$, $\hat{X}_{1,n}$ and $X^*_{1,n}$.

$$\|X_1-\hat{X}_1\| \le \|X_1-X^*_{1,n}\| \le \|X_1-\hat{X}_{1,n}\|$$
$$(2.8)\qquad \le \|X_1-\hat{X}_1\| + \|\sum_{m=n+1}^{\infty} A_m X_{-m}\| .$$

For the last inequality to hold it is necessary to assume that $\hat{X}_1$ has a mean convergent autoregressive representation and this will be assumed throughout this paper. Also, for the norm of the error of approximation of $\hat{X}_1$ by $X^*_{1,n}$ we have

$$(2.9)\qquad \|\hat{X}_1-X^*_{1,n}\| \le \|\hat{X}_{1,n}-X^*_{1,n}\| + \|\sum_{m=n+1}^{\infty} A_m X_{-m}\| .$$

Let σ denote the norm of the one-step ahead prediction error vector, that is $\sigma=\|X_1-\hat{X}_1\|^2=trG=tr\Phi_0^2$. It follows from (2.8) that

$$\sigma_n-\sigma \le \| X_1-\hat{X}_{1,n}\| -\sigma \le \| \sum_{m=n+1}^{\infty} A_m X_{-m}\| . \tag{2.10}$$

Thus the rate of convergence of $\sigma_n-\sigma$ is determined by that of $\| \sum_{n+1}^{\infty} A_m X_{-m}\|$. And the rate of convergence of the latter is determined by $q-\sigma_0/\sigma_n$ as it follows from Theorem 3.1 (b). The next section is devoted to the study of the rate of convergence of $q-\sigma_0/\sigma_n$, when the spectral density matrix f has sufficiently smooth entries. We note that for a univariate process, i.e. $q=1$, $\sigma_n-\sigma$ and $q-\sigma_0/\sigma_n$ have exactly the same rate of convergence (since $\sigma=\sigma_0$). But, the situation is completely different for $q > 1$.

3. Rates of Convergence of $\delta_n=q\sigma_n-\sigma_0$.

Since $\{\sigma_n\}$ is a bounded decreasing sequence it follows that the rate of convergence of $q\sigma_n-\sigma_0$ is the same as that of $q-\sigma_0/\sigma_n$. The next theorem which is a multivariate extension of Theorem 1 in Devinatz [3] plays an important role in estimating the rate of convergence of δ_n. Parts (b) and (c) of this theorem is used in finding rates of convergence for the linear predictor and interpolator of $\{X_n\}$.

Theorem 3.1

(a) If $f^{-1}\varepsilon L^1_{qxq}$ and H is any matrix-valued trigonometric polynomial of degree n with $H\ge\gamma I$, $0<\gamma<\infty$, then, $q-\sigma_0/\sigma_n\le \lambda^{-1}\int tr(f^{-1}-H)f(f^{-1}-H)^*dm$, where $\lambda>0$ is a constant.

(b) If $f\ge cI$, $0<c<\infty$, then,

$$q-\sigma_0/\sigma_n\ge c\sum_{m>n} trD_m D^*_m.$$

(c) If $cI\le f\le dI$, $0<c\le d<\infty$, and if $S_n=\sum_{k=0}^{n} D_k e^{ik\theta}$

with $trS_nS^*_n \le\gamma<\infty$ for all n, then

$$q-\sigma_0/\sigma_n\ge\left(\frac{c}{1+\sqrt{d\gamma}}\right)^2 \sum_{|m|>n} tr\, \hat{f}_m \hat{f}^*_m ,$$ where $\hat{f}_m$ is the m^{th} Fourier

coefficient f^{-1}.

Proof. (a) Let $f=\Phi\Phi^*$, $H=P^*P$, where Φ and P are the optimal factors of f and H respectively, cf. [8], [13, p. 504]. Denoting by Φ_0 and P_0 the constant terms of these H^2_{qxq} functions, we have

$$
\begin{aligned}
\int \operatorname{tr}(f^{-1}-H)f(f^{-1}-H)dm &= \int \operatorname{tr}(\Phi^{*-1}\Phi^{-1}-P^*P)\Phi\Phi^*(\Phi^{*-1}\Phi^{-1}-P^*P)dm \\
&= \int \operatorname{tr}(\Phi^{-1}P^{-1}-\Phi^*P^*)PP^*(P^{*-1}\Phi^{*-1}-P\Phi)dm \qquad (3.1)\\
&\geq \gamma\int \| P\Phi-(\Phi^*P^*)^{-1}\|^2_E\, dm.
\end{aligned}
$$

In the above argument we have used the fact that for compatible matrices A,B trAB=trBA, and that the eigenvalues of $(PP^*)(\theta)$ and $(P^*P)(\theta)$ are the same, and they are greater than γ.

Since $P\Phi\in H^2_{qxq}$, it follows that $P\Phi-P_0^{-1}\Phi_0^{-1}$ is orthogonal to $(\Phi^*P^*)^{-1}-P_0^{-1}\Phi_0^{-1}$. therefore,

$$
\begin{aligned}
\text{L H S of } (3.1) &\geq \gamma\int\|P_0^{-1}\Phi_0^{-1}-P\Phi\|^2_E dm=\gamma\int \operatorname{tr}(I-\Phi_0P_0P\Phi)(\Phi_0P_0^2\Phi_0)^{-1}(I-\Phi_0P_0P\Phi)^*dm \\
&\geq \gamma\lambda^{-1}\int\|I-\Phi_0P_0P\Phi\|^2_E dm,
\end{aligned}
$$

where λ is the largest eigenvalue of $\Phi_0P_0^2\Phi_0$, and the desired result follows from Lemma 2.1 (a), (b).

(b) From Corollary 2.2 we have

$$
\begin{aligned}
q-\sigma_0/\sigma_n &= \int \operatorname{tr}(\Phi^{-1}-\sqrt{\sigma_0}\, U_n)f(\Phi^{-1}-\sqrt{\sigma_0}\, U_n)^*dm \\
&\geq c\int\|\Phi^{-1}-\sqrt{\sigma_0}\, U_n\|^2_E dm \\
&\geq c\sum_{m=n+1}^{\infty} \operatorname{tr} D_mD_m^*,
\end{aligned}
$$

where the last inequality follows from the matricial Parseval's identity, cf. [15I, p. 121].

(c) By the matricial Parseval's identity we have

$$
\sum_{|m|>n} \operatorname{tr}\hat{f}_m\hat{f}^*_m \leq \int\|f^{-1}-S_n^*S_n\|^2_E\, dm \leq c^{-1}\int\|\Phi^{-1}-\Phi^*S_n^*S_n\|^2_E dm.
$$

Also,

$$\|\Phi^{-1}-\Phi^* S_n^* S_n\|_E \leq \|\Phi^{-1}-S_n\|_E + \|S_n-\Phi^* S_n^* S_n\|_E,$$

and

$$\|S_n-\Phi^* S_n^* S_n\|_E^2 = \mathrm{tr}(I-\Phi^* S_n^*)S_n S_n^*(I-\Phi^* S_n^*)^*$$

$$\leq (\mathrm{tr}\ S_n S_n^*)\,\mathrm{tr}(\Phi^{-1}-S_n)\Phi\Phi^*(\Phi^{-1}-S_n)$$

$$\leq \gamma d \|\Phi^{-1}-S_n\|_E^2.$$

Thus,

$$\|\Phi^{-1}-\Phi^* S_n^* S_n\|_E \leq (1+\sqrt{\gamma d})\|\Phi^{-1}-S_n\|_E.$$

and

$$\sum_{|m|>n} \mathrm{tr}\ \hat{f}_m \hat{f}_m^* \leq \left(\frac{1+\sqrt{\gamma d}}{c}\right)^2 \int \|I-\sqrt{\sigma_0}\, U_n \Phi\|_E^2 dm$$

$$= \left(\frac{1+\sqrt{\gamma d}}{c}\right)^2 (q-\sigma_0/\sigma_n),$$

by Lemma 2.1(b).

Q.E.D.

To find the rate of convergence of $\delta_n = q\sigma_n - \sigma_0$ when the entries of f are sufficiently smooth we need to have a matricial extension of the following important (univariate) theorem due to Grenander and Rosenblatt [4,5]: A necessary and sufficient condition that $\delta_n = \sigma_n - \sigma_0$ decreases at least exponentially to zero as n tends to infinity, is that $f(\theta)$ coincides in $[-\pi,\pi]$ almost everywhere with a function which is analytic for real θ and has no real zeros.

In the rest of this section we prove a matricial extension of this theorem. Our method of proof is similar to that of Devinatz [3]. However, in the proof of the necessity part given here, we do not use the properties of orthogonal polynomials with respect to f. Instead, we use a lemma due to Miamee [9] which provides a simpler and more direct proof of this part. In this section by the zeros of a matrix-valued function f we mean the zeros of det f.

The following theorem provides a matricial extension of the sufficiency part of the theorem of Grenander and Rosenblatt.

Theorem 3.2. Let $f \geq cI$, $0<c<\infty$, with entries that coincide in $[-\pi,\pi]$ almost everywhere with functions which are analytic for real θ, then

$$\delta_n = q\sigma_n - \sigma_0 = O(\rho^n), \quad 0 \leq \rho < 1.$$

Proof. Since $f \geq cI$, it follows that the entries of $f^{-1} = (f^{ij})$ are analytic in an open annulus $\{ Z \in \mathbb{C};\ \rho < |Z| < \frac{1}{\rho} \}$, $0 \leq \rho < 1$. By using an argument similar to that in [3, pp. 114-115] we get that $\hat{f}_k^{ij}$, the k^{th} Fourier coefficient of $\hat{f}^{ij}$ satisfies $f_k^{ij} = O(\rho^k)$ and that S_n, the symmetric partial sum of the Fourier series of f^{-1}, converges uniformly to f^{-1}. Thus, we have from Theorem 3.1 (a) that for n large

$$\delta_n = q\sigma_n - \sigma_0 = O\Big(\sum_{|k|>n} \operatorname{tr} \hat{f}_k \hat{f}_k^* \Big),$$

from which the result follows.

Q.E.D.

By using an argument similar to that in the proof of Theorem 3.2, and Theorem 2 [14] we obtain a slower rate of convergence for δ_n under much weaker differentiability requirement on the entries of f. It is interesting to compare this result with Theorem 3.1 and its special case in Baxter [1, p. 138].

Theorem 3.3. Let $f \geq cI$, $0<c<\infty$, and d be a positive integer.

(a) If all the entries of $f^{-1} = (f^{ij})$ have 2d-th derivative on $[-\pi,\pi]$, and the Fourier series of these 2d-th derivatives converge, then

$$\delta_n = q\sigma_n - \sigma_0 = 0(n^{-2d}).$$

(b) If all the entries of $f = (f_{ij})$ have 2d-th derivative on $[-\pi,\pi]$, and the Fourier series of these 2d-th derivatives are absolutely summable, then

$$\delta_n = q\sigma_n - \sigma_0 = 0(n^{-2d}).$$

Next, we prove a multivariate version of the necessity part of the theorem of Grenander and Rosenblatt. For a matrix $\Psi \varepsilon L^1_{qxq}$, we define a norm for Ψ by $\|\Psi\|_1 = \int \|\Psi\|_E dm$. The following lemma plays an important role in our proof of this result. For completeness we sketch the proof of this lemma [9].

Lemma 3.4 (Miamee). If $L^2(f) \subset L^1_{qxq}$, then there exists a positive constant M such that

$$\|\Psi\|_1 \le M \|\Psi\|_{L^2(f)} \text{, for all } \Psi \varepsilon L^2(f).$$

Proof. It is enough to show that the operator $T: L^2(f) \to L^1_{qxq}$ defined by $T(\Psi) = \Psi$ is closed, then the result follows from the closed graph theorem.

Q.E.D.

The assumption $L^2(f) \subset L^1_{qxq}$ is a natural one, since it guarantees that the Fourier coefficients of any function in the spectral domain is well-defined. Furthermore, it is not restrictive as the following simple lemma shows.

Lemma 3.5. $L^2(f) \subset L^1_{qxq}$ if $f^{-1} \varepsilon L^1_{qxq}$, and only if $(\det f)^{-\frac{1}{2q}} \varepsilon L^1$.

The next theorem, which is of independent interest, shows that if δ_n goes to zero sufficiently rapidly, then the entries of f^{-1} have summable Fourier series [3].

Theorem 3.6. If $\sum_{n=1}^{\infty} 2^n \delta^{\frac{1}{2}}_{2^n} < \infty$ and $L^2(f) \subset L^1_{qxq}$, then the entries of f^{-1} have summable Fourier series.

Proof. We have from Corollary 2.2, Lemma 3.5 and Cauchy-Schwartz inequality that

$$\begin{aligned}
(q-\sigma_0/\sigma_n)^{\frac{1}{2}} &= \left(\int tr(\Phi^{-1} - \sqrt{\sigma_0}\, U_n) f (\Phi^{-1} - \sqrt{\sigma_0}\, U_n)^* dm\right)^{\frac{1}{2}} \\
&\ge M^{-1} \int \|\Phi^{-1} - \sqrt{\sigma_0}\, U_n\|_E dm \\
&\ge (Mq)^{-1} \sum_{i,j=1}^{q} \int |\phi^{ij} - \sqrt{\sigma_0}\, U_{n,ij}| dm \\
&\ge (Mq)^{-1} \sum_{i,j=1}^{q} \sup_{\ell \ge 0} |(\phi^{ij} - \sqrt{\sigma_0}\, U_{n,ij})^{\wedge}(\ell)| \\
&\ge (Mq)^{-1} \sum_{i,j=1}^{q} |\hat{\phi}^{ij}(k)| ,
\end{aligned}$$

for $k \geq n+1$, where $U_{n,ij}$ denotes the ij-th element of U_n.

Thus, we have

$$\sum_{k=2^n+1}^{2^{n+1}} |\hat{\phi}_k^{ij}| \leq K\, 2^n\, \delta_{2^n}^{\frac{1}{2}}, \quad n=1,2,\ldots,i\ ,j=1,\ldots,q,$$

and summing both sides over n we get

$$\sum_{k=0}^{\infty} |\hat{\phi}_k^{ij}| < \infty, \quad i,j=1,2,\ldots,q,$$

which implies that the entries of Φ^{-1} have summable Fourier series. Hence, the same is true of $f^{-1}=\Phi^{*-1}\Phi^{-1}$, cf. [7, p. 31].

Q.E.D.

Now we can prove the multivariate version of the necessity part of the Grenander and Rosenblatt's theorem.

Theorem 3.7. If $\delta_n = O(\rho^n)$, $0 \leq \rho < 1$. Then f has no real zeros and all the entries of f are analytic on $[-\pi,\pi]$.

Proof. Since $\sum_{n=1}^{\infty} \delta_n < \infty$, it follows from Theorem 3.6 that $f \geq cI$ for some constant $0<c<\infty$. Thus, from Theorem 3.1 (b)

$$\sum_{k=n+1}^{\infty} \operatorname{tr} D_k D_k^* \leq K \rho^n,$$

and this implies that the entries of Φ^{-1} and Φ^{*-1} are analytic functions on $[\pi,\pi]$. Hence, the entries of $f^{-1}=\Phi^{*-1}\Phi^{-1}$ are analytic. Now, f^{-1} can have no real zeros since this would contradict the fact that f is in L^1_{qxq}. Thus, $f=(f^{-1})^{-1}$ is analytic with no real zeros.

Q.E.D.

4. Rates of Convergence of Finite Linear Predictors

In this section by using the results of the previous section we find rates of convergence for the finite linear predictors when the process $\{X_n\}$ has a smooth spectral density function. Here, we confine our attention to the case when the entries of f are analytic functions of θ. Similar results can be obtained when entries of f have properties like those stated in Theorem 3.3.

When $\hat{X}_1$ has a mean convergent autoregressive representation it follows from (2.8) that the finite linear least squares predictor of X_1, i.e. $X^*_{1,n}$ converges to $\hat{X}_1$ in the norm of M^q. According to (2.9) the rate of convergence in this case is dominated by $\|\sum_{n+1}^{\infty} A_m X_{-m}\|$ and $\|\hat{X}_{1,n} - X^*_{1,n}\|$. It is shown by Wiener and Masani [15 II] that the boundedness condition

$$cI \le f \le dI, \quad 0 < c \le d < \infty, \tag{4.1}$$

is sufficient for the existence of a mean convergent autoregressive representation of $\hat{X}_1$. For weaker conditions and more up to date results of this type, see [11, 12]. It follows from (4.1) that

$$\begin{aligned} \|\sum_{n+1}^{\infty} A_m X_{-m}\|^2 &\le d \int \|\sum_{n+1}^{\infty} A_m e^{im\theta}\|_E^2 dm(\theta) \\ &= d \sum_{n+1}^{\infty} \operatorname{tr} A_m A_m^* \qquad (4.2) \\ &\le d\,\sigma \sum_{n+1}^{\infty} \operatorname{tr} D_m D_m^*, \end{aligned}$$

and thus from Theorem 3.1 (b) we have

$$\|\sum_{n+1}^{\infty} A_m X_{-m}\| = O(\delta_n^{1/2}). \tag{4.3}$$

This combined with Theorem 3.2 and (2.8) gives the following

Theorem 4.1. Let $\{X_n\}$ be a q-variate WSSP with the spectral density matrix f satisfying (4.1) whose entries are analytic function of θ. Then, there exists a ρ, $0 \le \rho < 1$, such that

(a) $\|\hat{X}_1 - \hat{X}_{1,n}\| = O(\rho^n)$.

(b) $\sigma_n - \sigma = O(\rho^n)$.

(c) $\|X_1 - \hat{X}_{1,n}\| - \sigma = O(\rho^n)$.

(d) $\operatorname{tr} D_k D_k^* = O(\rho^n)$ and $\operatorname{tr} C_k C_k^* = O(\rho^n)$.

Next we concentrate on finding the rate of convergence of $\|\hat{X}_1-\hat{X}^*_{1,n}\|$. Because of (2.9) and (4.3) it is enough to find the rate of convergence of $\|X^*_{1,n}-\hat{X}_{1,n}\|$. For q=1, this has been found by Baxter [1] using a certain inequality and convergence equivalence involving Szegö's orthogonal polynomials. It seems very difficult to obtain matricial extension of these results. In the following we obtain the rate of convergence of $\|X_{1,n}-\hat{X}_{1,n}\|$ when $1\leq q<\infty$, without using the notion of orthogonal polynomials.

The rate of convergence of $\|X^*_{1,n}-\hat{X}_{1,n}\|$ is of fundamental importance in the statistical theory of multiple time series, particularly in studying the asymptotic properties of the autoregressive estimator of the spectrum, cf. [2,6]. This importance is revealed by the following simple fact: If $f\geq cI$, then

$$(4.5)\qquad \sum_{m=1}^{n} \|A_{nm}-C_0 D_m\|_E^2 = O(\|X^*_{1,n}-\hat{X}_{1,n}\|),$$

which provides a rate of convergence to zero for the entries of the matrices $A_{nm}-C_0 D_m$, n=1,2,...

In the following we find the rate of convergence of $\|X^*_{1,n}-\hat{X}_{1,n}\|$ by using a slightly different but equivalent characterization of $X^*_{1,n}$ than that introduced and used in Sections 2 and 3. This is based on identifying $X^*_{1,n}$ as the unique minimizer of the matricial functional $\psi(Y)=(X_1-Y,\ X_1-Y)$, $Y\in L=\overline{sp}\{X_{-k};\ 0\leq k\leq n-1\}$. We say that $Y_0\in L$ is a <u>minimizer</u> of $\psi(.)$ and write $\psi(Y_0)=\min_{Y\in L}\psi(Y)$ if

$$\psi(Y_0)\leq\psi(Y),\ \text{for all } Y\in L.$$

It is shown in [8, p. 354] that such a Y_0 exists and it is unique. This unique minimizer is $X^*_{1,n}$, indeed. Next we characterize the isomorph of $X^*_{1,n}$ in the spectral domain. Let

$$(4.6)\qquad G_n=\min \int PfP^*dm,$$

where the min is taken over all n^{th} degree matricial polynomial $P(\theta)=\sum_{m=0}^{n} P_m e^{im\theta}$ with $P_0=I$. If $\mathcal{U}_n(\theta)=\sum_{m=0}^{n} \mathcal{U}_{nm} e^{im\theta}$ is the minimizer of (4.6), then $\mathcal{U}_n$ can be characterized as the unique n^{th} degree polynomial with $\mathcal{U}_{n0}=I$ for which

$$\int \mathcal{U}_n f\, e^{im\theta}\, dm=0, \quad 1\leq m\leq n.$$

A different and more useful characterization of $\mathcal{U}_n$ is given by letting $V_n=G_n^{-1}\mathcal{U}_n$ and showing that

$$\min\int (I-P\Phi)(I-P\Phi)^*dm=$$

$$(4.7)\qquad \int (I-\Phi_0 V_n\Phi)(I-\Phi_0 V_n\Phi)^*dm=I-\Phi_0 G_n^{-1}\Phi_0=$$

$$\int (\Phi^{-1}-\Phi_0 V_n)f(\Phi^{-1}-\Phi_0 V_n)^*dm,$$

where the min is taken over all n^{th} degree matricial polynomials. It should be noted that the identities in (4.7) are essentially the matricial version of the results of Lemma 2.1 and Corollary 2.2. Let $S_n(\theta)=\sum_{m=0}^{n} D_m e^{im\theta}$, with these notations and the fact that $G_n V_{n0}=C_0 D_0=I$ we have

$$\|X^*_{1,n}-\hat{X}_{1,n}\| = \|\sum_{m=0}^{n}(G_n V_{nm}-C_0 D_m)X_{-m}\|$$

$$= \|\sum_{m=0}^{n}(G_n V_{nm}-GV_{nm}+GV_{nm}-GG^{-\frac{1}{2}}D_m)X_{-m}\|$$

$$\leq \|G_n-G\|_E \|\sum_{m=0}^{n} V_{nm}X_{-m}\| + \|G^{\frac{1}{2}}\|_E \|\Phi_0 V_n-S_n\| .$$

We have

$$G_n-G\leq tr(G_n-G).I=(\sigma_n-\sigma)I,$$

Therefore,

$$(4.9)\qquad \|G_n-G\|_E=O(\sigma_n-\sigma).$$

The rate of convergence of $\|\Phi_0 V_n - S_n\|$ can be obtained as follows: From (4.7) with $f \geq cI$ we have

$$c^{-\frac{1}{2}}[\operatorname{tr}(I - G^{\frac{1}{2}} G_n^{-1} G^{\frac{1}{2}})]^{\frac{1}{2}} \geq (\int \| \Phi^{-1} - \Phi_0 V_n \|_E^2 dm)^{\frac{1}{2}} \geq$$

$$(\int \| S_n - \Phi_0 V_n \|_E^2 dm)^{\frac{1}{2}} - (\sum_{m>n} \operatorname{tr} D_m D_m^*)^{\frac{1}{2}},$$

therefore,

(4.10) $$(\int \| S_n - \Phi_0 V_n \|_E^2 dm)^{\frac{1}{2}} \leq (\sum_{m>n} \operatorname{tr} D_m D_m^*)^{\frac{1}{2}} + c^{-\frac{1}{2}}[\operatorname{tr}(I - G^{\frac{1}{2}} G_n^{-1} G^{\frac{1}{2}})]^{\frac{1}{2}}.$$

From Theorem 3.1 (b) we have

(4.11) $$\sum_{m>n} \operatorname{tr} D_m D_m^* = O(\delta_n),$$

as for $\operatorname{tr}(I - G^{\frac{1}{2}} G_n^{-1} G^{\frac{1}{2}})$ we have from the simple identity

$$I - G^{\frac{1}{2}} G_n^{-1} G^{\frac{1}{2}} = G^{-\frac{1}{2}}(G_n - G) G_n^{-1} G^{\frac{1}{2}},$$

that

$$\operatorname{tr}(I - G^{\frac{1}{2}} G_n^{-1} G^{\frac{1}{2}}) = \operatorname{tr} G_n^{-1}(G_n - G) \leq (\operatorname{tr} G_n^{-2})^{\frac{1}{2}} [\operatorname{tr}(G_n - G)^2]^{\frac{1}{2}} \leq K \| G_n - G \|_E.$$

Thus, from (4.9) we have

(4.12) $$\operatorname{tr}(I - G^{\frac{1}{2}} G_n^{-1} G^{\frac{1}{2}}) = O(\sigma_n - \sigma).$$

By combining these estimates with (2.9), Theorems 3.2 and 4.1 (b) we get the following

Theorem 4.2. Under the conditions of Theorem 4.1 we have

(a) $$\sum_{m=0}^{n} \| D_k - \Phi_0 G_n^{-1} A_{nm} \|_E^2 = \int \| S_n - \Phi_0 V_n \|_E^2 dm = O(\rho^n).$$

(b) $$\| \hat{X}_{1,n} - X_{1,n}^* \| = O(\rho^n).$$

(c) $$\| \hat{X}_1 - X_{1,n}^* \| = O(\rho^n).$$

Theorem 3.1 (c) can be used to find the rate of convergence for the <u>finite linear interpolator</u> of $\{X_n\}$. For this let us assume that all the values of $\{X_n\}$ are known, except for the value X_0. Then, under the assumption that $f^{-1} \varepsilon L^1_{qxq}$, it is known [10] that the linear least squares interpolator of X_0, denoted by $\check{X}_0$, can formally be written as

$$\check{X}_0 \sim -\hat{f}_0^{-1} \sum_{k \neq 0} \hat{f}_{-k} X_k \quad , \tag{4.13}$$

where $\hat{f}_k$ is the k^{th} Fourier coefficient of the matrix f^{-1}. Sufficient conditions for the mean convergence of the series in (4.13) are given in [10]. In particular, it follows from Theorem 7 [10] that the boundedness condition (4.1) guarantees the convergence of this series.

Now one can form a finite interpolator for X_0, denoted by $\check{X}_{0,n}$ and defined by $\check{X}_{0,n} = -\hat{f}_0^{-1} \sum_{\substack{k=-n \\ k \neq 0}}^{n} \hat{f}_{-k} X_k$. The error in norm of this approximation is given by

$$\|\check{X}_0 - \check{X}_{0,n}\|^2 \leq d \, \|\hat{f}_0^{-1}\|^2 \int_E \Big\| \sum_{|k|>n} \hat{f}_{-k} e^{ik\theta} \Big\|^2 dm =$$

$$= d \, \|\hat{f}_0^{-1}\|^2 \sum_{|k|>n} \operatorname{tr} \hat{f}_k \hat{f}_k^* .$$

Thus, from Theorem 3.1 (c) we have

$$\|\check{X}_0 - \check{X}_{0,n}\| = O(\delta_n^{1/2}),$$

and combining this with Theorem 3.2 we get the following

<u>Theorem 4.3</u>. Under the conditions of Theorem 4.1 we have

$$\|\check{X}_0 - \check{X}_{0,n}\| = O(\rho^n), \quad 0 \leq \rho < 1.$$

An exponential rate of convergence for the finite linear interpolator of a univariate stationary process, when its past and future are at positive angle [10,11], is obtained by Salehi [13]. However, in his work the finite linear interpolator is obtained by using the Von Neumann's alternating projections, as a result it is less feasible for computation since it involves the one-sided innovation process $\{Y_n\}$ and the coefficients of the optimal factor of the density. An interesting by product of Salehi's approach is the identification of ρ in Theorem 4.3 with the cosine of the angle between the past and future of the process $\{X_n\}$. In view of this, one may raise the following natural and useful question: Is it possible to identify the constant ρ appearing in this paper (say, in Theorem 3.2) with the cosine of the angle between the past and future of the process $\{X_n\}$? If the answer is positive then in an obvious manner, the results of this paper confirm the intuitively appealing fact that the closer this angle is to $\pi/2$ the faster the finite predictors will converge to the infinite predictor $\hat{X}_1$.

Another problem which is worthy of study in connection with this work is the infinite-dimensional generalization of the results presented here. Presence of q in Theorem 3.1 and δ_n shows that the method of proof used here is intrinsically finite-dimensional. It is of interest to know, e.g. the analogues of Theorem 3.1 and δ_n in the infinite dimensional case.

REFERENCES

1. BAXTER, G. (1962). An asymptotic result for the finite predictor. Math. Scand. 10, 137-144.

2. BERK, K.N. (1974). Consistent autoregressive spectral estimates. Ann. Statist. 2, 489-502.

3. DEVINATZ, A. (1964). Asymptotic estimates for the finite predictor. Math. Scand. 15, 111-120.

4. GRENANDER, V. and ROSENBLATT, M. (1954). An extension of a theorem of G. Szegö and its application to the study of Stochastic Processes, Trans. Amer. Math. Soc. 76, 112-126.

5. GRENANDER, U. and SZEGÕ, G. (1958). Toeplitz Forms and their Applications, University of California. Press, Berkeley.

6. HANNAN, E.J. and HONG-ZHI, A. and ZHAO-GUO, C. (1982). Autocorrelation, autoregression and autoregressive approximation. Ann. Statist., 10, 926-936.

7. KATZNELSON, Y. (1968). An Introduction to Harmonic Analysis. Dover, New York.

8. MASANI, P. (1966). Recent trends in multivariate prediction theory. Multvariate Analysis (ed. P. R. Krishnaiah) 351-382. Academic Press, New York.

9. MIAMEE, A.G. (1984). On the angle between past and future and prediction theory of stationary stochastic processes. (Preprint).

10. POURAHMADI, M. (1984). On the mean convergence of the best linear interpolator of multivariate stationary stochastic processes. Ann. Probab. 12, 609-614.

11. POURAHMADI, M. (1984). A matricial extension of the Helson-Szegő Theorem and its application in multivariate prediction. J. of Multivariate Analysis, to appear.

12. POURAHMADI, M. (1984). The Helson-Sarason-Szegö Theorem and the Abel summability of the series for the predictor. Proc. of the Amer. Math. Soc., 91, 306-308.

13. SALEHI, H. (1974). On bilateral linear interpolator for minimal stationary stochastic processes. SIAM J. Appl. Math. 26, 502-507.

14. TONG, J. (1984). A note on the degree of approximation of Fourier series. Arch. Math. 42, 157-158.

15. WIENER, N. and MASANI, P. (1957). The prediction theory of multivariate stochastic processes. Part I. Acta Math. 98, 111-150; Part II. Acta Math. 99 (1958) 93-137.

END

FILMED

10-84

DTIC